PRZEGLĄD GEOGRAFICZNY
T. L, Z. 2, 1978

ALICJA Z. SZAJNOWSKA

# Zastosowanie entropijnych miar w analizie przestrzennej

*The application of entropic measures in a spatial analysis*

Zarys treści. Autorka w wielkim skrócie informuje o metodologicznych możliwościach i analitycznych korzyściach stosowania entropijnych miar do zagadnień zróżnicowania rozmieszczenia przestrzennego zjawisk i procesów geograficznych.

Zjawiska i struktury geograficzne przedstawiają sobą najczęściej złożone wieloczynnikowe systemy związane z przestrzennym rozmieszczeniem. Takie ujęcie zróżnicowania rozmieszczenia przestrzennego zawiera tezę konieczności i celowości stochastycznego spojrzenia na zdarzenia i procesy geograficzne.

Systematyczne studia (ze stochastycznego punktu widzenia) nad tematyką zróżnicowania rozmieszczenia przestrzennego zapoczątkował B. L. Gurewicz (1966, 1967). Utrzymuje on, że ze zróżnicowaniem geograficznym związane są w sposób naturalny entropijne miary jednorodności (homogeniczności) i zróżnicowania przestrzennego, które mogą służyć jako podstawa konkretnej analizy rozmieszczenia i są niezbędne przy tworzeniu ogólnej teorii zróżnicowania przestrzennego. Za miarę jednorodności i zróżnicowania przestrzennego przyjmuje się miarę entropii. Jest ona jednym z podstawowych pojęć cybernetycznych i pozwala na analizę zachowania się układów stochastycznych.

Według Małego Słownika Cybernetycznego (1973) entropia, to miara jednorodności i stopnia zróżnicowania elementów lub stanów znajdujących się w pewnym zbiorze przeliczalnym, które traktowane są przy określaniu ich możliwej wartości jako zmienne losowe.

Entropię danej zmiennej losowej można obliczyć, znając charakterystyki stochastyczne tej zmiennej. Entropia określonej zmiennej losowej ma następujące wartości

1. entropia jest funkcją ciągłą względem wszystkich swych charakterystyk statystycznych, rozumianych jako zmienne. Dla ciągłej zmiennej losowej taką charakterystyką jest gęstość prawdopodobieństwa jej realizacji f(x), natomiast dla dyskretnej zmiennej losowej — prawdopodobieństwo wystąpienia $i$-tej realizacji $p_i$,

2. entropia jest tym większa przy ustalonym zakresie zmienności, im bardziej rozkład prawdopodobieństwa zmiennej losowej jest zbliżony do rozkładu równomiernego,

3. dla zbioru niezależnych zmiennych losowych entropia jest sumą entropii jego podzbiorów,

4. entropia jest równa zeru dla zmiennej losowej, której zbiór wartości jest równy jedności.

Jedyną funkcją spełniającą cztery wymienione warunki jednocześnie jest funkcja logarytmiczna. Podstawa logarytmu określa bezwymiarową jednostkę pomiaru entropii.

Entropia dyskretnej (nieciągłej) zmiennej losowej Xd odznacza się ciągiem rozkładów $p_i = P\{Xd = x_i\}$ dla $i = 1, 2, ..., n$ i jest określona wzorem:

$$H_{(Xd)} = -\sum_{i=1}^{N} p_i \log_a p_i = \sum_{i=1}^{N} p_i \log \frac{1}{p_i}, \tag{1}$$

gdzie:

$p_i$ — prawdopodobieństwo $i$-tej realizacji dyskretnej zmiennej losowej,

$x_i$ — $i$-ta realizacja zmiennej losowej $Xd$.

Jeżeli logarytm ma podstawę a=2, to miara entropii jest wyrażona w bitach:

$$0 \leqslant H_{Xd} \leqslant 1.$$

## Dotychczasowe zastosowania miar entropii

Zdaniem R. Lee (1974) dotychczasowe próby adaptacji i zastosowania entropii w badaniach przestrzennych można skrótowo przedstawić w dwóch grupach tematycznych:

I. liczbowa analiza map, w której entropia służy do wyrażenia miary różnorodności próbki z mapy,

II. metody maksymalnej (maksimum) entropii — jako miary różnorodności przestrzennej zjawisk i procesów stochastycznych.

I. Entropia w liczbowych analizach map jest przedstawiona jako wyrażenie opisowej miary różnorodności lub stopnia przypadkowości próbki z danej mapy.

Ju. Miedwiedkow (1976a) do liczbowej analizy map użył wzoru Shannona (1), gdzie zero oznacza brak entropii, tzn. punkty pobranej próbki z mapy są jednolite.

Entropia jako miara podziału powierzchni zmienności cech została zastosowana do:

1. zagadnień przestrzennego podziału przez B. L. Gurewicza (1969),

2. porównania alternatywnych podziałów każdej zbiorowości przez L. Ju. Nutenko (1970) i M. Battego (1972),

3. pomiaru wpływu środowiska naturalnego na produkcję rolną przez B. Marchand (1972),

4. porównania stopnia zróżnicowania rozmieszczeń miast wielkości różnej wielkości przez G. P. Chapman (1970),

5. przedstawienia przestrzennej koncentracji zatrudnienia w przemyśle przez C. G. Garrisona i A. S. Paulsona (1970).

Są to (w większości przytoczonych przypadków) teoretyczne rozważania i dyskusje na temat celowości adaptacji miar entropii do przedstawienia zjawisk i procesów przestrzennych albo niewielkie próby zastosowania, bez podania końcowych wyników lub wniosków.

Nieco inny wariant wzoru (1) zastosował G. P. Chapman (1970, 1973) i M. Batty (1972), a potem H. Theil (1967) do oznaczania zjawisk rozproszenia. Takie kryterium klasyfikacji danego zjawiska nie zawiera więcej niż jedną charakterystykę. Stąd zasygnalizowane modele bardzo często różnią się jedynie specyfiką odzwierciedlanych zjawisk i procesów. W związku z tym termin „model" w tych przypadkach jest błędnie użyty. Jest to miara opisowa, podobnie jak w metodzie najbliższego sąsiada, w analizie równań czy w ilorazie lokalizacji.

T. K. Semple i R. G. Golledge (1970) w swoich studiach porównawczych (z zastosowań) usiłowali wprowadzić element czasu, tzn. obliczone punkty entropijnego modelu traktowali ponadczasowo. Dyskusyjne innowacje metodyczne zawierają również opracowania: Ju. Miedwiedkowa (1967b), T. K. Sempla (1973) i G. P. Chapmana (1973).

II. Metody maksymalnej entropii dają bardziej ogólną (teoretyczną) podstawę metodyczną niż zastosowania liczbowych analiz mapy. Poza tym technika maksimum entropii pozwala na dedukcyjne analizy najbardziej zawiłych podziałów „przestrzeni geograficznych". Miara entropii maksimum najpierw była traktowana jako wektor określony na podstawie hierarchii np. porządku pokoleń czy też skali regionu. Początkowe pomysły na tym polu były rozwijane niezależnie przez wielu badaczy.

A. G. Wilson (1967) podał cały wachlarz możliwości zastosowań metody maksymalnej entropii, przede wszystkim do określenia przestrzennego rozkładu wyjazdów wycieczkowych za miasto. Inne zastosowania ukierunkowano następująco:

— modele regionalnych przepływów artykułów żywnościowych (A. G. Wilson, 1970a, 1970b),
— model lokacyjne (A. G. Wilson, 1970c),
— przestrzenna struktura miasta (P. L. Fano, 1969),
— dystrybucja dochodu regionalnego (M. J. H. Magride, 1969),
— teoria centralnego placu (B. Marchand, 1972),
— przestrzenna geometria regionu (M. Batty, 1972),
— Model Markowa do przedstawienia współzależności przestrzennych (S. G. Tomlin, 1969), który rozszerzył R. B. Ginsberg (1971), wprowadzając element czasu, co w semi-modelu Markowa powoduje, że prawdopodobieństwa przejścia są funkcją czasu.

B. J. L. Berry i P. J. Schwind (1969) uważają, że w sytuacji poważnych luk w literaturze metodologicznej na temat wzajemnych oddziaływań przestrzennych, miara entropii jest jedyną metodą zobrazowania przepływów migracyjnych. Próbę takiego zastosowania przedstawili przejrzyście i systematycznie. Po długim wstępie wprowadzającym w koncepcję zastosowań miary entropii, a przede wszystkim po jej matematycznej prezentacji, autorzy podają przykład ilustracji przepływów migracyjnych przy pomocy macierzy entropii.

Poszczególne miary entropii są zawarte w macierzy $P$, gdzie $p_{ij}$ — przemieszczenia z $i$-tego do $j$-tego obszaru SEA's (State Economic Areas), Macierz entropii P:

$$H(p_{ij})=\sum_{i}^{N}\sum_{j}^{N}p_{ij}\log\frac{1}{p_{ij}} \quad \text{(wzajemne przemieszczenia)} \tag{2}$$

$$H(p_i)=\sum_{i}^{N}p_i\log\frac{1}{p_i} \quad \text{(wiersze)} \tag{3}$$

$$H(p_i)=\sum_{j}^{N}p_j\log\frac{1}{p_j} \quad \text{(kolumny)} \tag{4}$$

Tabela 1

Teoretyczne i empiryczne przepływy migracyjne między SEA's w stanie Iowa

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | teoretyczne | 5034 | 1063 | 1446 | 1146 | 885 |
| | empiryczne | 5202 | 1967 | 862 | 1024 | 1413 |
| 2 | 5034 | | 1373 | 2752 | 1716 | 1232 |
| | 4698 | | 2382 | 4539 | 2465 | 2301 |
| 3 | 1063 | 1373 | | 907 | 2312 | 1373 |
| | 2249 | 3614 | | 665 | 2301 | 2488 |
| 4 | 1446 | 2752 | 907 | | 1195 | 1111 |
| | 665 | 3748 | 367 | | 1331 | 3197 |
| 5 | 1146 | 1716 | 2312 | 1195 | | 1320 |
| | 539 | 2858 | 3576 | 1425 | | 3711 |
| 6 | 885 | 1232 | 1373 | 1111 | 1320 | |
| | 560 | 1343 | 1227 | 2127 | 2416 | |

*Źródło*: B. J. L. Berry i P. J. Schwind (1969)

Tabela 1 zawiera liczbowe przemieszczenia ludności między 6 SEA's w stanie Iowa.

W tabeli 2 odnotowano porównawcze wyniki dwóch modeli migracji: modelu grawitacji, modelu entropii.

Tabela 2

Miary entropii dla macierzy migracji

| Model entropii | Wiersze | Kolumny | Wzajemne przemieszczenia |
|---|---|---|---|
| Maksimum ($p_{ij}$) | 1.079 | 1.079 | 2.120 |
| Empiryczne ($m_{ij}$) | 1.037 | 1.008 | 1.897 |
| Model grawitacji | 1.025 | 1.025 | 1.972 |

*Źródło*: B. J. L. Berry i P. J. Schwind (1969).

Dla macierzy migracji 6 × 6 maksimum entropii dla wierszy i kolumn wynosi po 1.075, a wzajemne przemieszczenia — 2.120. Entropia dla empirycznych przemieszczeń — 1.037 wiersze, 1.008 kolumny. Można to porównać z modelem grawitacji, który daje po 1.025 dla wierszy i kolumn i 1.975 dla wzajemnych przemieszczeń.

Modele Ju. Miedwiedkowa (1967a) i A. G. Wilsona (1967, 1970) oparte na formule Shannona (1) stanowią istotę wszystkich entropijnych modeli przestrzennych. Wobec powyższego, przykład zastosowania miar entropii przez M. G. Sonis (1968) do przedstawienia przemieszczeń ludności USA między 4 rejonami może być podstawową literaturą pozwalającą na zapoznanie się z techniką entropijnych miar jednorodności.

Jak już znaznaczyłam, zdaniem wielu badaczy, entropijne miary jednorodności i zróżnicowania przestrzennego przyczyniają się do tworzenia i rozszerzania teorii przemieszczeń. Ze stochastycznego punktu widzenia każde przemieszczenie ludności można interpretować jako rozmieszczenie, gdyż rozmieszczenie jest rezultatem pewnego procesu przemieszczeń. W takiej sytuacji można entropijne miary jednorodności rozmieszczenia przenieść na przypadek przemieszczeń. Aby tak interpretować, trzeba dokonać pewnej operacji na zbiorach punktów (regionów) migracji. Przyjmujemy 2 zbiory: $\Omega_1$ i $\Omega_2$. Iloczynem tych zbiorów jest zbiór $\Omega = \Omega_1 \cap \Omega_2$ zawierający wszystkie możliwe pary punktów ($M_1$, $M_2$), gdzie punkt $M_1$ należy do zbioru $\Omega_1$, a punkt $M_2$ — do zbioru $\Omega_2$:

$$\Omega = \Omega_1 \cap \Omega_2 = \{(M_1, M_2):(M_1 \in \Omega_1, M_2 \in \Omega_2\}.$$

Gdy rozpatrujemy przemieszczenia, dzielimy obszar $\Omega$ na $n$ obszarów, a następnie bierzemy pod uwagę migracje z $i$-tego do $j$-tego obszaru. Tę parę obszarów traktujemy jako jeden obiekt, któremu przypisujemy wagę $m_{ij}$ równą liczbie ludności migracyjnej z i-tego do j-tego obszaru. W ten sposób przemieszczenia stają się rozmieszczeniem (na podstawie iloczynu kartezjańskiego), którego elementami są pary obszarów z odpowiadającą każdej parze wagą $m_{ij}$.

Fakt, że przemieszczenia można interpretować jako rozmieszczenie, pozwala wykorzystać entropijną miarę jednorodności rozmieszczenia wyrażoną wzorem (1), a przede wszystkim przyczynia się do rozszerzenia samej teorii zróżnicowania. Wzór (1) można zapisać:

$$H = -\sum_{ij}^{N} \log \frac{m_{ij}}{M} \log m_{ij} = \log M - \frac{1}{M} \sum_{i,j=1}^{N} m_{ij} \log m_{ij}, \tag{5}$$

gdzie logarytmy mają podstawę 2, zatem miara jednorodności określona w bitach. Jak widać, miara jednorodności H uwzględnia nie tylko różnice pomiędzy wielkościami $m_{ij}$, ale także i same wielkości $m_{ij}$ i ich sumę M. W ten sposób otrzymuje się ogólną charakterystykę procesu przemieszczenia z punktu widzenia jednorodności.

Tak pojętą miarę jednorodności zastosował M. G. Sonis (1968) do przedstawienia przemieszczeń ludności USA w 4 regionach: północno-wschodni, północ regionu centralnego, południowy i zachodni (tab. 3).

Tabela 3 jest macierzą kwadratową zawierającą elementy $m_{ij}$ (i, j=1, 2, 3, 4), które przedstawiają sobą liczebność osób przemieszczających się z $i$-tego regionu do $j$-tego regionu. Suma osób znajdujących się na przekątnej jest liczbą osób zmieniających ewentualnie adres wewnątrz regionu bez wyjazdów do innego regionu.

W tabeli 4 przedstawiono wartości miar jednorodności migracji dla ludności USA między czterema regionami w okresie 3 lat, obliczone według

Tabela 3

Rozmieszczenie migrantów w 4 regionach USA

| Regiony | Północno-wchodni | Północ regionu centralnego | Południowy | Zachodni |
|---|---|---|---|---|
| Północnowschodni | 1548 | 153 | 253 | 111 |
| Północ regionu centralnego | 128 | 2043 | 383 | 160 |
| Południowy | 299 | 422 | 3305 | 315 |
| Zachodni | 184 | 410 | 507 | 2021 |

*Źródło*: M. G. Sonis (1968).

Tabela 4

Miary jednorodności

| | Lata | Ludność USA (4 regiony) |
|---|---|---|
| H (w bitach) | 1963 | 3.19 |
| | 1964 | 3.17 |
| | 1965 | 3.16 |

wzoru (5). Odczytuje się z niej wyraźną stabilność miary jednorodności migracji ludności USA w badanym okresie.

W konkluzji wymienieni autorzy stwierdzają, że cybernetyczne ujęcie zagadnienia przemieszczeń migracyjnych, a w szczególności model migracji pozwala na nowe i dodatkowe możliwości analizy procesów migracyjnych. B. J. L. Berry i P. J. Schwind (1969), Ju. W. Miedwiedkow (1970) G. P. Chapman (1970) przedstawili nie tylko formułę miary entropii i jej wprowadzenie w kontekst dynamiki modelu Markowa, lecz także dokładnie opisaną macierz zastosowanej eneropii. Zdaniem R. Lee (1974), te próby zastosowań nie zostały w pełni podsumowane i ocenione.

## BIBLIOGRAFIA

Batty M., 1972. *Entropy and spatial geometry.* "Area", 4, 230—236.

Berry B. J. L., P. J. Schwind, 1969. *Information and entropy in migrant flows.* "Geographical Analysis", 1, 5—14.

Chapman G. P., 1970. *The applications of information theory to the analysis of population distributions in space.* "Economic Geography", 46, 317—331.

Chapman G. P., 1973. *The spatial organization of the population of the United States and England and Wales.* „Economic Geography". 49, 325—343.

Fano P. L., 1969. *Organization, city size distributions, and central places.* "Papers of the Regional Science Association", 22, 29—38.

Garrison C. B., A. S. Paulson, 1970. *An entropy measure of the geographic concentration of economic activity.* "Economic Geography", 49, 319—324.

Ginsberg R. B., 1971. *Semi-Markow processes and mobility.* "Journal of Mathematical and Sociology", 1, 233—262.

Gurewicz B. L., 1969. *Measures of Feature-based and areal differentiatins and* *their use in city services.* "Soviet Geography: Review and Translation", 10, 383—386.

G u r e w i c z B. L., Ju. G. S a u s z k i n. 1966. *Matiematiczeskij mietod w gieografii.* "Wiestnik Moskowskogo Uniwersitieta", Sierija V — Gieografija, 1, 3—27.

G u r e w i c z B. L., 1967. *Płotnost nasielenija goroda i płotnost wierojatnosti słuczajnoj wieliczіny.* "Wiestnik Moskowskogo Uniwersitieta", Sierija V — Gieografija, 1, 15—21.

L e e R., 1974. *Entropy models in spatial analysis.* "Discussion Paper", 15, 1—53.

Mały Słownik Cybernetyczny. 1973, Wiedza Powszechna, 103—107.

M a r c h a n d B., 1972. *Information theory and geography.* "Geographical Analysis", 4, 234—257.

M i e d w i e d k o w Ju., 1967a. *The regular component in settlement patterns as shown on a map.* "Soviet Geography: Review and Translation", 8, 150—168.

M i e d w i e d k o w Ju., 1967b. *Concept of entropy in settlement pattern analysis.* "Papers of the Regional Science Association", 18, 165—168.

M i e d w i e d k o w Ju., 1970. *Entropy: an assessment of potentialities in geography.* "Economic Geography", 46, 306—316.

M o g r i d g e M. J. H., 1969. *Some factors influencing the income distribution of households within a city region* (W:) *J. J.* S c o t t. *Studies in Regional Science,* London, Pion, 117—141.

N u t e n k o L. J., 1970. *An information — theory approach to the partitioning of an area.* "Soviet Gegoraphy: Review and Translation", 8, 540—544.

S e m p l e R. K., 1973. *Recent trands in the spatial concetraion of corporate headguaters.* "Economic Geography", 49, 309—318.

S e m p l e R. K., R. G. C o l l e d g e, 1970. *An analysis of entropy changes in a settlement pattern over time.* "Economic Geography", 46, 157—160.

S o n i s M. G., 1968. *Znaczenije entropijnych mier odnorodnosti dla analiza pierieraspriedielienij nasiclenja* (W:) *Matiematika w ekonomiczeskoj gieografii.* „Woprosy Gieografii", 77, 44—63.

T h e i l H., 1967. *Economics and information theory.* Amsterdam, North-Holland.

W i l s o n A. G., 1967. *A statistical theory of spatial distribution models.* "Transportation Research", 1, 253—269.

W i l s o n A. G., 1970a. *Entropy in urban and regional modelling.* London, Pion.

W i l s o n A. G., 1970b. *Inter-regional commodity flows: entropy maximising approaches.* "Geographical Analysis", 2, 255—282.

W i l s o n A. G., 1970c. *Disaggregating elementary residential location models.* "Papers, Regional Science Association", 24, 103—125.

АЛИЦИЯ З. ШАЙНОВСКА

## ПРИМЕНЕНИЕ ЭНТРОПИЙНЫХ МЕР В ТЕРРИТОРИАЛЬНОМ АНАЛИЗЕ

Автор на основании доступной литературы сообщает о методологических возможностях и аналитических выгодах применения энтропийных мер для измерения дифференциации территориального размещения географических явлений и процессов. Систематическое изучение тематики территориальной дифференциации было начато Б. Л. Гуревичем (1966, 1967), который утверждает, что с географической дифференциацией естественным образом связаны энтропийные меры гомогенности и территориальной дифференциации. Они могут служить основой для конкретного анализа размещения и необходимы при создании общей

теории территориальной дифференциации. В качестве меры гомогенности и территориальной дифференциации принята мера энтропии на базе формулы Шеннона:

$$H_{xd} = \sum_{i=1}^{N} p_i \log_a p_i,$$

где:

$P_i$ — вероятность $i$-той реализации дискретной случайной переменной,
$X_i$ — $i$-тая реализация случайной переменной $Xd$.

Если у логарифма основание a=2, то мера энтропии выражена в битах.

По мнению Р. Ли (1974), проводимые до сих пор попытки адаптировать и применить энтропию в территориальных исследованиях можно кратко представить в двух тематических группах:

1. Численный анализ карт, в котором энтропия служит для выражения меры разнородности выборки с карты;

2. Методы максимальной (максимум) энтропии как меры территориальной разнорадности стохастических явлений и процессов.

Пер. Б. Миховского

ALICJA Z. SZAJNOWSKA

## THE APPLICATION OF ENTROPIC MEASURES IN A SPATIAL ANALYSIS

The author utilized available literature to survey the methodological possibilities and analytical advantages of applying entropic measures in studies concerned with the differentiation of the spatial distribution of geographical phenomena and processes.

A methodical investigation of this subject was started by B. L. Gurewicz (1966, 1967). This author believes that entropic measures of homogeneity and spatial differentiation, which can be used in an analysis of the distribution and which are necessary when a general theory of spatial differentiation is being worked out, are connected in a natural way with the geographical differentiation. Shannon's formula:

$$H_{(Xd)} = - \sum_{i=1}^{N} P_i \log_a P_i$$

where:

$P_i$ = probability of the i-th realization of the discrete random variable,

$x_i$ — i-th realization of the random variable Xd usnally serves as a basis for measuring homogeneity and spatial differentiation.

If the basis of the logarithm is a=3, the measure of entropy is expressed in bits.

R. Lee (1974) suggests that the attempts, made so far, to adapt and apply entropy in spatial research can be grouped as follows: 1. Numerical analyses of maps, in which entropy can serve as a means to express the measure of heterogeneity of the sample from the map. 2. The method of the maximal (maximum) entropy, as a measure of spatial heterogeneity of stochastic phenomena and processes.

Translated by *Halina Dzierzanowska*